

Abstract of the Disclosure

SYSTEM AND METHOD

FOR

AUTOMATICALLY CLASSIFYING TEXT

A method is provided for automatically classifying text into categories. In operation, a plurality of tokens or features are manually or automatically associated with each category. A weight is then coupled to each feature, wherein the weight indicates a degree of association between the feature and the category. Next, a document is parsed into a plurality of unique tokens with associated counts, wherein the counts are indicative of the number of times the feature appears in the document. A category score representative of a sum of products of each feature count in the document times the corresponding feature weight in the category for each document is then computed. Next, the category scores are sorted by perspective, and a document is classified into a particular category, provided the category score exceeds a predetermined threshold.